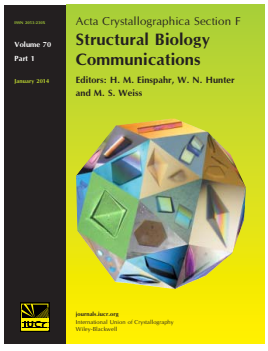
Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

Structure of transportin SR2, a karyopherin involved in human disease, in complex with Ran

Vicky G. Tsirkone, Katrien G. Beutels, Jonas Demeulemeester, Zeger Debyser, Frauke Christ and Sergei V. Strelkov

Acta Cryst. (2014). F**70**, 723–729



Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

Structure of transportin SR2, a karyopherin involved in human disease, in complex with Ran

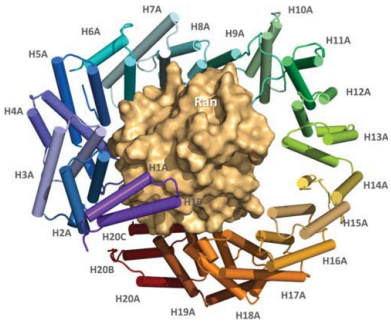
Vicky G. Tsirkone,[a] Katrien G. Beutels,[a] Jonas Demeulemeester,[b] Zeger Debyser,[b] Frauke Christ[b] and Sergei V. Strelkov[a]*

[a]Laboratory for Biocrystallography, Department of Pharmaceutical and Pharmacological Sciences, KU Leuven, Herestraat 49 bus 822, 3000 Leuven, Belgium, and [b]Laboratory of Molecular Virology and Gene Therapy, Department of Pharmaceutical and Pharmacological Sciences, KU Leuven, Herestraat 49 bus 822, 3000 Leuven, Belgium

Correspondence e-mail: sergei.strelkov@pharm.kuleuven.be

Received 21 March 2014
Accepted 28 April 2014

PDB reference: transportin SR2, 4ol0



Transportin SR2 (TRN-SR2) is a β-type karyopherin responsible for the nuclear import of specific cargoes, including serine/arginine-rich splicing factors. The protein has been implicated in a variety of human diseases, including HIV infection, primary biliary cirrhosis and limb-girdle muscular dystrophy 1F. Towards understanding its molecular mechanism, a 2.9 Å resolution crystal structure of human TRN-SR2 complexed with the small GTPase Ran has been determined. TRN-SR2 is composed of 20 α-helical HEAT repeats forming a solenoid-like fold. The first nine repeats form a 'cradle' for the binding of RanGTP, revealing similarities but also differences with respect to the related importin 13 complex.

1. Introduction

Compartmentalization in the eukaryotic cell allows tight regulation of biological functions such as transcription and translation, but it also necessitates active transport of complex macromolecules in and out of the nucleus *via* the nuclear pore complex (NPC; Hoelz *et al.*, 2011). Such transport relies on soluble factors (importins and exportins) belonging to the karyopherin-β family (Stewart, 2007; Chook & Süel, 2011). The karyopherins shuttle their cargoes between the cytoplasm and the nucleus *via* a GTP-dependent process which relies on a small GTPase called Ran (Ras-related nuclear protein). Specifically, nuclear import is driven by the gradient of the GTP-bound form of Ran, which is enriched in the nucleus. After binding the cargo in the cytoplasm, the importin docks to the cytoplasmatic side of the NPC and is then transferred to the other side, driven by transient interactions with nucleoporins (D'Angelo & Hetzer, 2008). Once in the nucleus, the importin–cargo complex dissociates owing to the competitive binding of RanGTP to the importin, followed by the return of the importin–RanGTP complex to the cytoplasm and GTP hydrolysis (Lee *et al.*, 2005).

Here, we focus on human transportin SR2 (TRN-SR2), a β-type importin (Lai *et al.*, 2000). Both TRN-SR2 and its splice variant TRN-SR1 are encoded by the *tnpo3* gene (Yun *et al.*, 2003). A characteristic feature of TRN-SR2 is its ability to import serine/arginine-rich proteins (SR proteins), in particular splicing factors. Like other family members, TRN-SR2 is predicted to consist of 20 consecutive α-hairpin motifs (two antiparallel α-helices joined by a short linker) known as HEAT repeats (Andrade & Bork, 1995). The stacked HEAT repeats are generally believed to create a structure with considerable plasticity, which provides adaptive binding of different cargoes and regulatory proteins (Cook *et al.*, 2007). While forming a single continuous physical domain, the HEAT-based karyopherins include distinct and sometimes overlapping functional regions. In particular, their N-terminal part binds RanGTP, whereas the C-terminal part is generally responsible for cargo binding (Lai *et al.*, 2000).

Importantly, TRN-SR2 has been implicated in a variety of human diseases, including human immunodeficiency virus (HIV) infection (Brass *et al.*, 2008; Christ *et al.*, 2008; König *et al.*, 2008), primary biliary cirrhosis (PBC; Hirschfield *et al.*, 2010) and limb-girdle muscular dystrophy 1F (LGMD1F; Melià *et al.*, 2013; Torella *et al.*, 2013). The role of TRN-SR2 in the HIV life cycle has been well

studied. Yeast two-hybrid screening initially identified TRN-SR2 as a direct binding partner of HIV integrase (Christ *et al.*, 2008), a finding that was later confirmed by other approaches (Krishnan *et al.*, 2010; Larue *et al.*, 2012). Using fluorescent techniques, it has been demonstrated that TRN-SR2 is directly involved in the nuclear import of the HIV pre-integration complex, leading to the establishment of infection (Christ *et al.*, 2008). Next, hereditary limb-girdle muscular dystrophy 1F was linked to a single nucleotide deletion in the stop codon of the *tnpo3* gene. This mutation was shown to result in an extended protein of either 938 (form *A*) or 1018 (form *B*) amino acids, compared with the native TRN-SR2 of 923 residues. The mutated protein accumulates in the outer rim of the nuclear membrane, suggesting disturbed nuclear transport (Melià *et al.*, 2013; Torella *et al.*, 2013). Given both the fundamental aspect of the TRN-SR2-dependent nuclear import and the above-mentioned disease context, it is not surprising that this protein has recently gained interest and has been actively studied at the biochemical and cellular levels (Christ *et al.*, 2008; De Houwer *et al.*, 2012; Larue *et al.*, 2012; Taltynov *et al.*, 2013). As the next major step, here we present the atomic structure of human TRN-SR2, which was made possible by resolving its complex with Ran at 2.9 Å resolution using X-ray crystallography.

Table 1
Crystallographic data-collection and refinement statistics.

	Overall	Inner shell	Outer shell
Data collection			
Space group	$P2_12_12_1$		
Unit-cell parameters (Å)	a = 74.2, b = 110.0, c = 148.5		
Resolution range (Å)	44–2.9	44–13	3.0–2.9
No. of unique reflections	27515	348	2035
Multiplicity	4.1	3.5	4.1
Completeness (%)	99.6	95.3	99.9
$\langle I/\sigma(I)\rangle$	6.9	15.5	1.6
$CC_{1/2}$†‡	0.990	0.986	0.516
R_{merge}‡	0.146	0.052	1.00
R_{meas}‡	0.168	0.061	1.15
Wilson B factor (Å^2)	60.7		
Refinement			
R_{work}	0.222		
R_{free}§	0.276		
No. of non-H atoms	8625		
R.m.s.d., bonds (Å)	0.003		
R.m.s.d., angles (°)	0.79		
Ramachandran favoured/outliers (%)	94.3/0.2		
MolProbity score	1.8		
Average B factor (Å^2)	73.5		

† Half-set correlation of intensities. ‡ As output by *AIMLESS* (Evans, 2006). § As calculated using a 'free' set of 1381 reflections (5%).

2. Materials and methods

Both full-length human TRN-SR2 and full-length human RAN mutant Q69L (which is practically incapable of GTP hydrolysis) were overexpressed in *Escherichia coli* BL21 (DE3) Rosetta pLysS cells using the pETHSUMO vector (Weeks *et al.*, 2007). The expressed products were fusions containing an N-terminal 6×His tag, the small ubiquitin-like modifier (SUMO) protein, a SUMO hydrolase cleavage site and the protein sequence of interest. After digesting the vector with *Hin*dIII and *Kpn*I, the sequences encoding *Homo sapiens* TRN-SR2 (residues 3–923) and Ran^{Q69L} (residues 1–216) were inserted using the In-Fusion system (Clontech).

Bacterial cultures were grown at 24°C until an $OD_{600\ nm}$ of 4.0 was reached, and then for a further 24 h at 18°C. The cells expressing TRN-SR2 and Ran^{Q69L} were pelleted by centrifugation and resuspended in buffer *A* (20 m*M* Tris pH 7.5, 200 m*M* NaCl, 15 m*M* imidazole, 5 m*M* β-mercaptoethanol) and buffer *B* (50 m*M* Tris pH 7.5, 150 m*M* NaCl, 4 m*M* $MgCl_2$, 15 m*M* imidazole, 10% glycerol, 2.5 m*M* β-mercaptoethanol), respectively. After the addition of 5 µl Benzonase per 100 ml, the cells were sonicated on ice for 2 min with 20% amplitude and 30/10 s on/off pulses using a Branson Digital Sonifier. The procedure was repeated twice with 15 min in between. The lysates were cleared at 18 000*g* for 30 min at 4°C and were then applied onto an Ni^{2+}–NTA (GE Healthcare) column using buffers *A* and *B* for TRN-SR2 and Ran^{Q69L}, respectively, followed by elution with the corresponding buffer supplemented with 500 m*M* imidazole. The tags were then cleaved off by overnight incubation with SUMO hydrolase at a 1:500 molar ratio at 4°C in buffers *A* and *B*, respectively. This was followed by a second pass through the Ni^{2+}–NTA column, whereby the flowthrough containing the cleaved protein was collected while the uncleaved protein and the protease, both containing the His tag, were retained on the column. For TRN-SR2, the buffer was subsequently exchanged to 20 m*M* Tris pH 7.5, 50 m*M* NaCl, 5 m*M* DTT using an Ultra-15 ultrafiltration device (Amicon), which was followed by ion-exchange chromatography on a 5 ml HiTrap Q HP column (GE Healthcare) using a linear gradient of 50 m*M* to 1 *M* NaCl. To load Ran with GTP, the protein was incubated for 30 min at room temperature with a fivefold molar excess of GTP and 20 m*M* EDTA. The loading was stopped by the addition of 40 m*M* $MgCl_2$ followed by dialysis against 50 m*M* Tris, 150 m*M* NaCl, 5 m*M* DTT pH 7.5. For complex formation, TRN-SR2 and Ran^{Q69L}GTP were mixed in a 1:2 molar ratio and incubated on ice for 2 h, followed by size-exclusion chromatography on a HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare) in 20 m*M* Tris pH 7.5, 50 m*M* NaCl, 5 m*M* DTT. Prior to crystallization, the complex was concentrated to 9.8 mg ml^{-1}.

Initial screening for crystallization of the TRN-SR2–Ran^{Q69L}GTP complex was performed using commercial screens and a Mosquito crystallization robot (TTP Labtech). Small needles could be obtained in 400 nl sitting drops using The Protein Complex Suite (Qiagen) condition 45 [100 m*M* MES pH 6.5, 5%(*v*/*v*) MPD, 15%(*w*/*v*) PEG 6000] at 20°C. Subsequently, larger crystals could be obtained in 3 µl hanging drops equilibrated against an equivalent reservoir solution but with a slightly higher pH (6.7). After several rounds of microseeding, the best crystal reached dimensions of 150 × 30 × 20 µm in 10 d. X-ray diffraction data were collected on the PROXIMA 2A beamline (Synchrotron Soleil, Saint-Aubin, France) using a microfocus 10 × 5 µm beam with 0.979 Å wavelength. The crystals were briefly soaked in mother liquor supplemented with 30% MPD and flash-cooled in liquid nitrogen. To limit radiation damage to the thin crystal, five different spots on the crystal were exposed; each was used to collect a 20° rotation wedge. Data were processed with *XDS* (Kabsch, 2010) and scaled with *AIMLESS* (Evans, 2006). The crystals contained one protein complex per asymmetric unit (Table 1).

Initially, we failed to phase the data by molecular replacement using as a search model our homology-modelled structure of the TRN-SR2–Ran^{Q69L}GTP complex (Taltynov *et al.*, 2013), as well as any available experimental structures of related importins or their complexes with Ran, including the complex of a closely related human importin 13 (Imp13) with yeast RanGTP (PDB entry 2x19; Bono *et al.*, 2010). In retrospect, this difficulty could be explained by significant differences in the overall conformation. However, it was possible to obtain the correct molecular-replacement solution, albeit with a marginal contrast, by searching with only the RanGTP from the latter structure using *MOLREP* (Vagin & Teplyakov, 2010). This can be rationalized by the remarkably high conservation between human and yeast Ran (90% sequence identity), even though this

protein makes up only 15% of the total complex mass. With the Ran molecule in place, a further *MOLREP* search could position a segment of the Imp13 structure from the complex containing HEAT repeats 1–5. With these two elements in place, another *MOLREP* run could correctly position a segment containing repeats 6–9. Such an assembled partial structure yielded a crystallographic free *R* factor of 0.54, which could be improved to 0.51 by automated refinement in *REFMAC* (Murshudov *et al.*, 2011), but only if the 'jelly-body' restraints were enabled. At this point, HEAT repeats 10–13 of Imp13 could be approximately placed into the electron-density map, followed by further refinement and placement of additional repeats until the model was complete. After multiple cycles of manual rebuilding using *Coot* (Emsley *et al.*, 2010) and refinement using *PHENIX* (Adams *et al.*, 2010), a final model with good quality was obtained (Table 1). Atomic coordinates and structure factors have been deposited in the Protein Data Bank (http://www.pdb.org) under accession code 4ol0. Figures were prepared with *PyMOL* (http://www.pymol.org/). The protein complexes were analyzed with the *PISA* algorithm (Krissinel & Henrick, 2007) using an online server (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html).

3. Results and discussion

Previously, we have studied the detailed biochemical properties of human TRN-SR2, and in particular characterized its complex with the catalytically inactive Q69L mutant of human Ran (Taltynov *et al.*, 2013). For the current work, we have changed the overexpression system for both proteins, which yielded highly pure tag-free samples. As a result, after forming the TRN-SR2–Ran^{Q69L}GTP complex, single crystals could be grown. Despite their small size, the use of a microfocus synchrotron beam and an optimized data-collection strategy allowed us to measure a complete 2.9 Å diffraction data set (Table 1). Phasing of the data only became possible *via* a carefully chosen molecular-replacement protocol, whereby the RanGTP molecule and a number of the N-terminal HEAT repeats of TRN-SR2 could be positioned. This was followed by gradual building of the remaining structure. Of note, we were also able to obtain large (>100 µm in each direction) crystals of human TRN-SR2 alone, but despite extensive optimization efforts these crystals diffracted X-rays to only about 6 Å resolution (O. Taltynov, V. G. Tsirkone and S. V. Strelkov, unpublished results).

The asymmetric unit of the TRN-SR2–Ran^{Q69L}GTP crystals contains a single copy of the complex. The entire TRN-SR2 molecule is well defined in the electron-density map, with the exception of three linkers located between HEAT repeats 5 and 6 (residues 216–220), the two helices of HEAT 14 (residues 598–605) and the two helices of HEAT 20 (residues 884–888). The C-terminal section (residues 180–216) of Ran^{Q69L} is also not visible, which is in line with earlier predictions that this part is disordered (Vetter *et al.*, 1999; Nilsson *et al.*, 2002). Previously, C-terminally truncated Ran had been used for the crystallization of the Imp13–Ran complex (Bono *et al.*, 2010).

Our crystal structure confirms that TRN-SR2 is a solenoid-type protein consisting of 20 HEAT repeats, corroborating this hallmark feature of the karyopherin-β family (Andrade & Bork, 1995). Each of the repeats is principally composed of two antiparallel α-helices connected through a loop, resulting in an 'α-hairpin'. The first α-helix of each repeat (called helix *A*) is always located on the outside of the solenoid, while the second α-helix (*B*) is on the inside (Fig. 1). The axes of helices *A* and *B*, although principally antiparallel, usually form a small crossing angle (about −15°, left-handed). Exceptionally, the C-terminal HEAT repeat 20 capping the open edge consists of three helices.

Neighbouring α-hairpins corresponding to HEAT repeats 2–9 are roughly parallel to each other (while the very first repeat is significantly rotated). Consequently, the overall conformation of the N-terminal part of TRN-SR2 corresponds to a sector of a straight solenoid (Fig. 1*a*). The same is valid for the C-terminal part (HEAT repeats 10–20). However, there is a pronounced rotation of HEAT 10 with respect to HEAT 9, which are connected by a particularly short linker. As a result, the overall shape of TRN-SR2 is that of a right-handed screw rather than of a straight solenoid (Fig. 1*b*).

The core of the Ran molecule is an SH3 domain fold consisting of six β-strands. This core is supplemented by three α-helices and two surface loops referred to as Switch I (residues 37–45) and Switch II

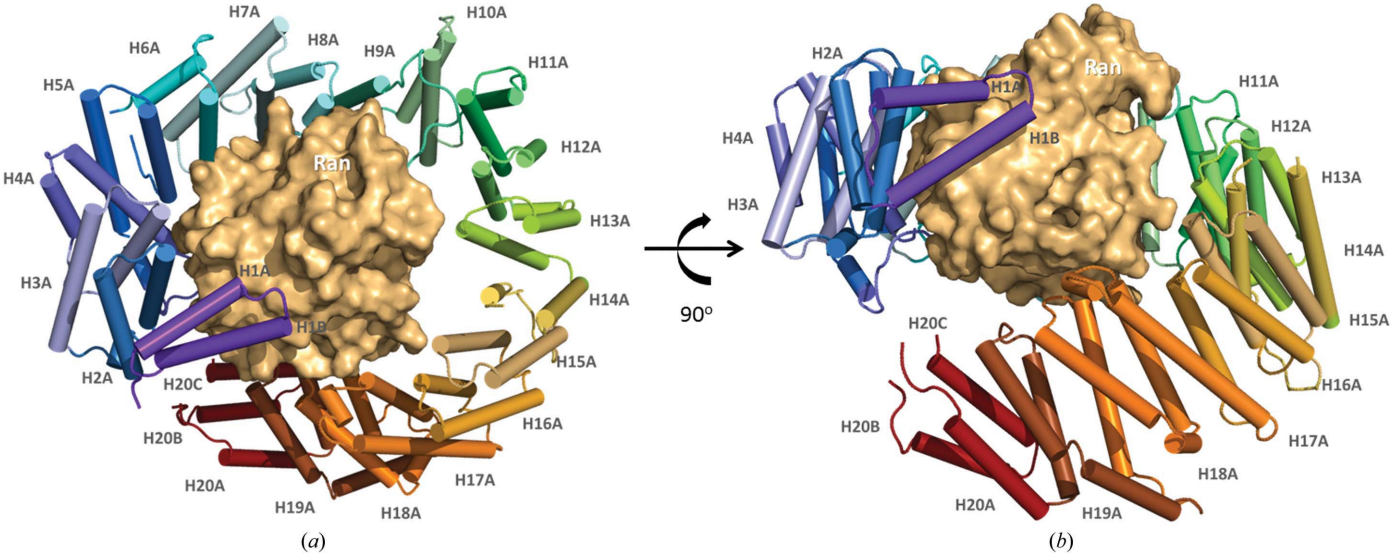
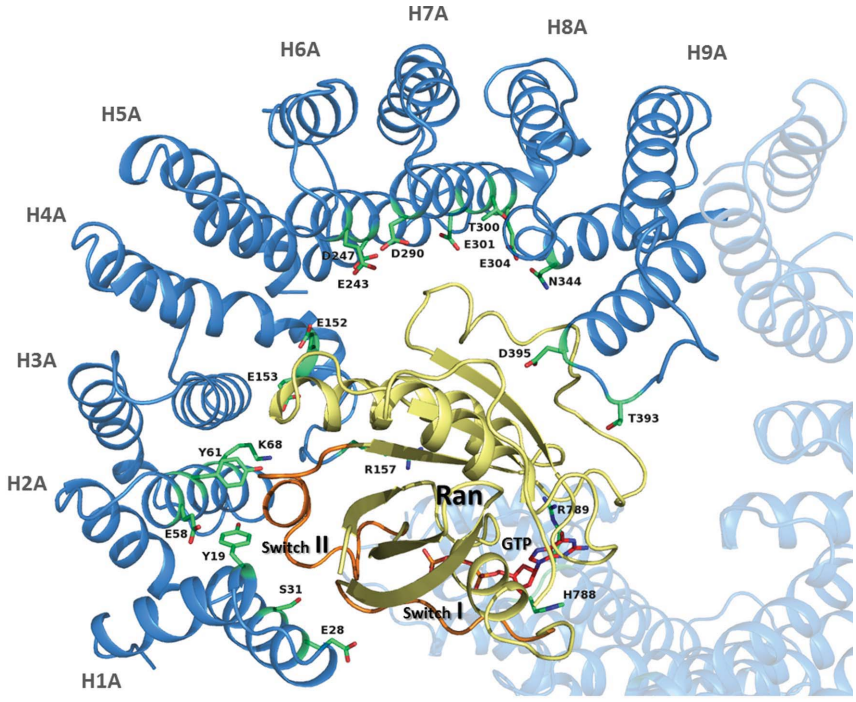


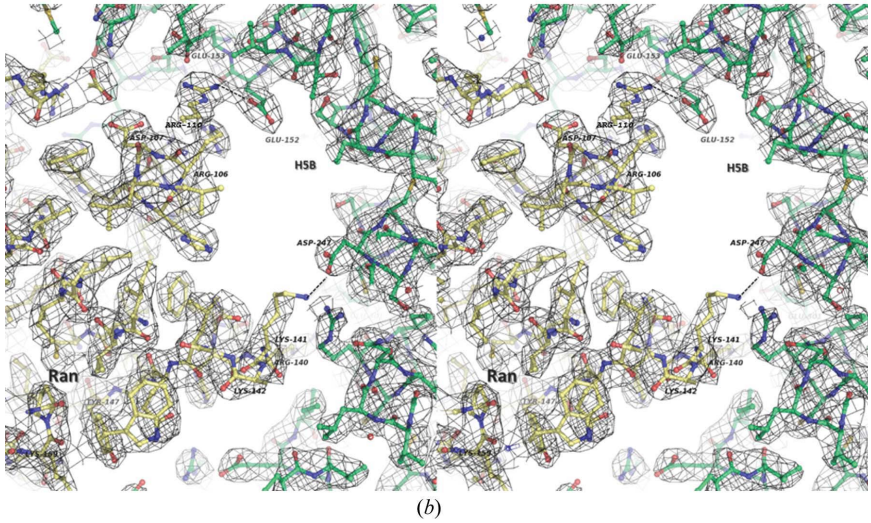
Figure 1
(*a*, *b*) Overall view of the TRN-SR2–Ran^{Q69L}GTP complex in two perpendicular orientations. Each HEAT of TRN-SR2 is shown in a different colour, with α-helices as cylinders. Ran is represented by a sand-coloured surface. Both helices *A* and *B* of HEAT 1 are labelled. The three helices *A*, *B* and *C* of HEAT 20 are also indicated. For all other repeats, only the outer helix *A* is labelled for clarity.

(residues 69–85) (Fig. 2). Both loops are known to undergo extensive conformational changes upon GTP binding (Stewart *et al.*, 1998).

The binding of Ran to TRN-SR2 requires GTP and results in a stable complex with an apparent K_d of 4.7 $\pm$ 1.2 n*M*, as we have shown *in vitro* using the AlphaScreen protein–protein interaction assay (Taltynov *et al.*, 2013). Indeed, our crystal structure reveals that formation of the complex buries a very considerable part (22.9%, 2081 Å^2) of the total surface area of RanGTP. The complex is stabilized in a synergistic manner by both hydrophobic interactions and polar contacts (Fig. 3). The latter interactions are represented by specific hydrogen bonds (to side groups or the main chain) and salt bridges across the interface (Figs. 2 and 4). The N-terminal part of TRN-SR2 clearly presents the specific site of RanGTP binding. Indeed, HEAT repeats 1–9 make a 'cradle' roughly matching the size of the globular Ran protein, albeit slightly larger. Correspondingly, the binding of Ran mostly involves the *B* (inner) helices of HEAT repeats 1–4 and 7–9, with some additional contacts made by repeats 5 and 6 (Figs. 2 and 3). On the Ran side, the exposed Switch II loop is involved in important interactions with HEAT 1 and HEAT 2, while Switch I makes further contacts with both HEAT 1 and HEAT 18 (Fig. 2). In addition, the N-terminal tips of helices *B* from repeats 15, 17 and 18 also make some contacts with the bound Ran (Fig. 3). It is likely that these latter interactions are not essential for RanGTP binding but rather occur simply because of the curved shape of the TRN-SR2 molecule. At the same time, the expected site of cargo binding is located in the C-terminal part of TRN-SR2 (Lai *et al.*,



(*a*)



(*b*)

Figure 2
Molecular detail of the interaction between TRN-SR2 and RanGTP. (*a*) Contacts of the main Ran-binding domain (HEAT repeats 1–9, intense blue). TRN-SR2 residues forming salt bridges and hydrogen bonds to Ran are shown in green with side chains as sticks. The Ran molecule is shown as a yellow ribbon, with the Switch I and Switch II regions highlighted in orange. The bound GTP is shown in red. (*b*) A close-up view of some interacting residues of TRN (green) and Ran (yellow) and the corresponding weighted $2|F_o| - |F_c|$ electron-density map at the 1σ level, presented in stereo. The salt bridges Glu152/Glu153(TRN)–Arg110(Ran) and Asp247(TRN)–Lys141(Ran) are shown with dashed lines.

2000). Since cargo release inside the nucleus is coupled to RanGTP binding, there must also be an interplay between the two processes at the structural level. Such interplay is facilitated by the solenoid shape of the TRN-SR2 molecule, since the two binding sites are located in proximity and may be sterically overlapping. Furthermore, the HEAT-based structure is clearly prone to conformational plasticity, as observed before for various crystallographic complexes of other importins (Lee *et al.*, 2006; Forwood *et al.*, 2010). This plasticity is likely to play an important role, whereby a particular binding partner could come in contact with just the N-terminal part or the C-terminal part, or with both of them as in the crystal structure presented here.

Finally, we have compared the structure of the TRN-SR2–Ran^{Q69L}GTP complex with the corresponding complex of Imp13, a β-karyopherin with 22.6% amino-acid identity and 38% similarity to

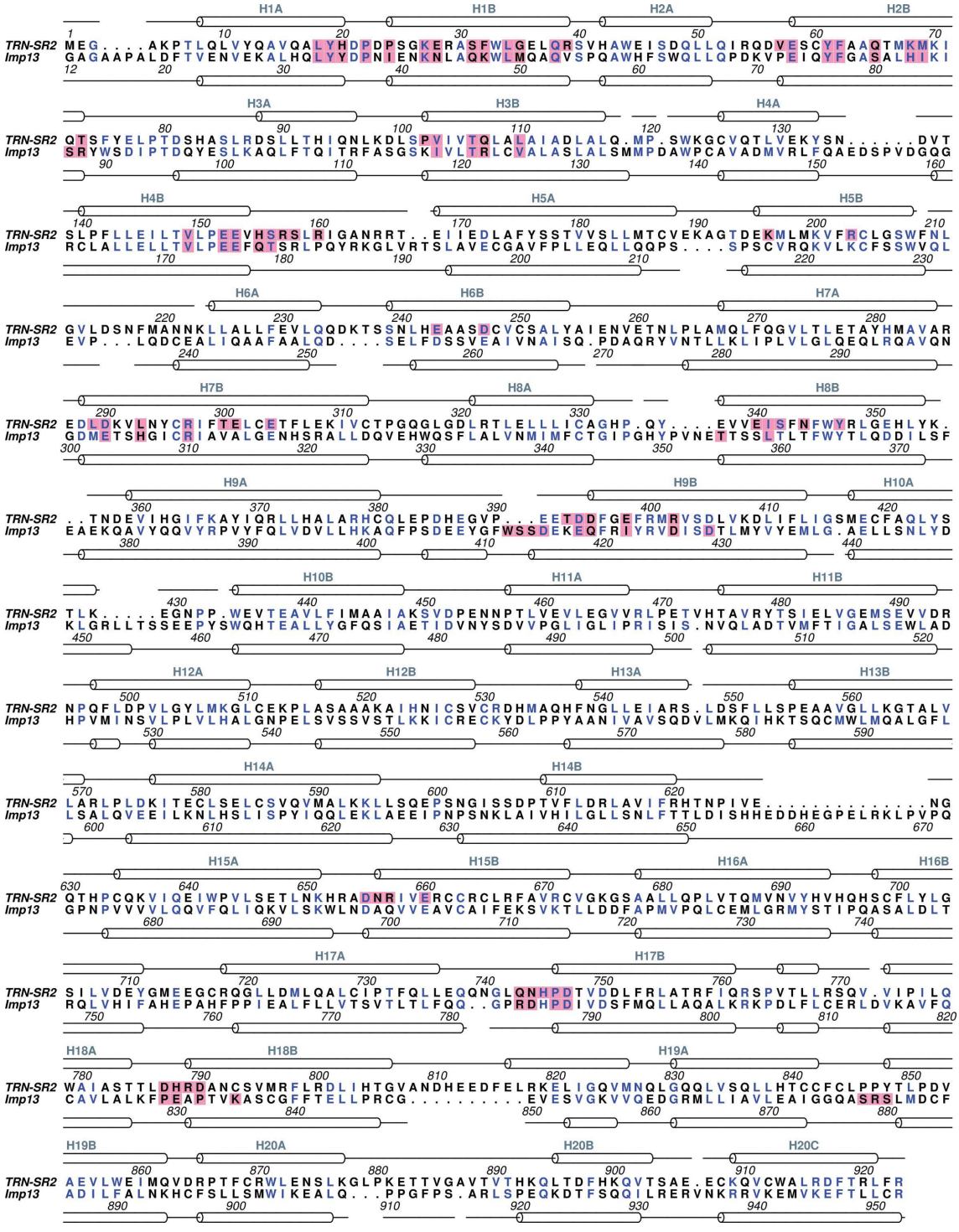


Figure 3
Structure-based sequence alignment of TRN-SR2 and Imp13 based on their respective complexes with RanGTP. The conserved residues (identical or similar type) are shown in blue. Magenta highlighting marks the residues that come into contact with RanGTP in the TRN-SR2–Ran^{Q69L}GTP structure (this work) and the Imp13 complex with yeast Ran (Bono *et al.*, 2010; PDB entry 2x19), as determined using *PISA* (Krissinel & Henrick, 2007). This includes van der Waals contacts, salt bridges and hydrogen bonds. The secondary-structure elements for TRN-SR2 and Imp13 are indicated above and below the alignment, respectively.

TRN-SR2 (Bono *et al.*, 2010). The latter complex includes Ran from the yeast *Saccharomyces cerevisiae* rather than human Ran (90% sequence identity). In line with the considerable homology between TRN-SR2 and Imp13, the overall HEAT-based structure of both proteins is similar (Fig. 5). However, the Imp13 solenoid wraps around the Ran globule in a somewhat less tight way than TRN-SR2, so that HEAT repeats 7 and 8 of Imp13 make only a few contacts with Ran, in contrast to our TRN-SR2 complex (Fig. 3). Moreover, there is some conservation but also considerable differences in the specific interactions (hydrogen bonding and salt bridges) that RanGTP makes with either of the two proteins. In particular, superposition of the two complexes reveals that as few as only four salt bridges (involving three Ran residues Asp77, Arg110 and Arg166) and three hydrogen bonds (involving Ran residues 47, 78 and 79) are structurally preserved (Fig. 4). Of note, the Ran-binding domain of these two importins (HEAT repeats 1–9) is somewhat more conserved (amino-acid sequence identity 26.0%) than the rest of the molecule (sequence identity of 20.1%).

Repeat	TRN-SR2		Ran		Imp13	Repeat
1	Tyr19	***	**Gly78**	***	Tyr34	1
1	Glu28	***	**Val47**	***	Asn43	1
1	Ser31	***	**Tyr79**	***	Gln46	1
			Val45	***	Gln46	1
2	Glu58	***	**Gln/Asn82**	***	Tyr35	1
2	Tyr61	***	Gly78			
2	Try61	***		***	Ser80	2
2	Lys68	+++	**Asp77**	+++	His83	2
				+++	Arg122	3
			Glu70	+++	Arg88	2
			Lys71	***		2
4	Glu152	***	Arg106			
			Asp107	+++	Arg122	3
4	Glu152	+++	**Arg110**	+++	Glu175	4
4	Glu153	+++		+++	Glu176	4
4	Arg157	+++	Asp18			
6	Asp247	+++	Lys141			
6	Glu243	+++				
7	Asp290	+++	Lys142			
7	Thr300	***				
7	Glu301	+++	Arg140			
7	Glu304	+++				
8	Asn344	***	Lys134			
9	Thr393	***	Lys159			
9	Asp395	+++	**Arg166**	+++	Asp415	9
9	Asp395	***	Tyr147			
15	Arg657	***	Glu34			
18	His788	***	**Lys37**	+++	Asp788	17
			Tyr39	***	Glu830	18
18	Arg789	***	Thr93			

Figure 4
Hydrogen bonds (***) and salt bridges (+++) formed between RanGTP and TRN-SR2 (left) and Imp13 (right) in their respective complexes. The interfaces were analyzed with *PISA* (Krissinel & Henrick, 2007). The interactions involving structurally equivalent pairs of residues in the two complexes are shaded grey. Ran residue 82 is Gln in human Ran (used for the TRN-SR2 complex) and Asn in yeast Ran (used for the Imp13 complex). Ran residues involved in the interaction in both complexes are shown in bold.

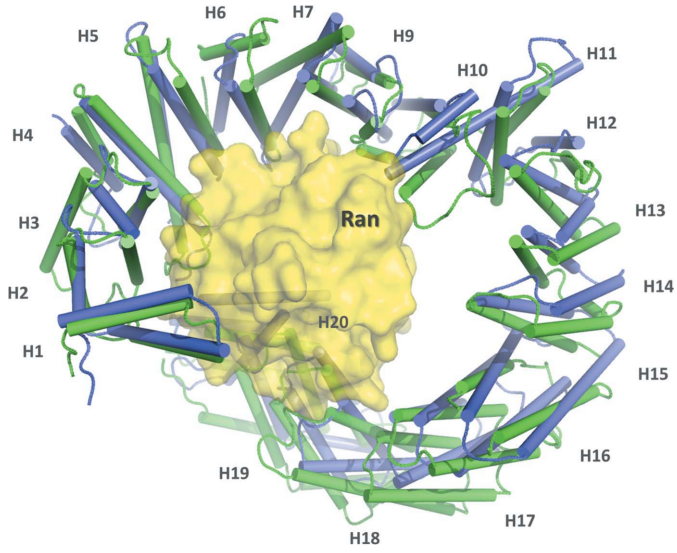


Figure 5
Comparison of the TRN-SR2–Ran^{Q69L}GTP complex and the Imp13 complex with yeast Ran (PDB entry 2x19), with the HEAT repeats indicated as H1–H20. The two complexes were superimposed by the Ran molecule (yellow semi-transparent surface). Overall, TRN-SR2 (green) and Imp13 (violet) wrap around Ran in a very similar manner, but some of the α-helices (cylinders) are considerably shifted.

As this paper was being written up, Maertens *et al.* (2014) reported the crystal structures of human TRN-SR2 on its own and in two complexes, including the Ran^{Q69L}GTP complex (PDB entry 4c0q). Compared with our structure, the latter structure was determined in a distinct space group *P*1 with two complexes per asymmetric unit at 3.42 Å resolution. Together with a better resolution, our structure additionally reveals several surface loops (residues 110–194, 354–362, 625–633 and 672–676). The overall conformation of TRN-SR2 in the two crystal structures is very similar, especially for the N-terminal part, which shows an identical mode of RanGTP binding. For a detailed comparison of the two structures, we have aligned them by HEAT repeats 1–15. Such a superposition yields an r.m.s. deviation of only 1.2 Å for all C$^{\alpha}$ atoms of HEATs 1–15. However, the C-terminal parts of the TRN-SR2 molecule (repeats 16–20) in the two structures gradually deviate more significantly. This appears to be owing to a different packing in the two crystals and to intrinsic flexibility of the TRN-SR2 solenoid. The largest differences in the C$^{\alpha}$ positions (up to 7 Å) are within the last two HEATs (19 and 20).

Determination of the atomic structure of TRN-SR2 will provide a basis for a fundamentally better understanding of its functioning at the molecular level. As outlined in §1, besides its involvement in other human diseases, TRN-SR2 plays a key role in HIV infection. Indeed, the nuclear import of the HIV pre-integration complex can be seen as the bottleneck for HIV replication, as it is an essential prerequisite for the integration of the viral genome into the host cell, which ultimately leads to persistent infection. We believe that targeting the interaction between TRN-SR2 and HIV integrase may be a promising avenue for new drug development, complementing existing therapies.

We thank Stephen Weeks for helpful advice on cloning and protein expression, Stephanie De Houwer for discussions and Maria-Despoina Charavgi for valuable advice on crystallography. This work was supported by the Research Foundation Flanders (FWO) grant G0665.12 (to SVS and FC). Research at the Laboratory of Molecular

Virology and Gene Therapy is also funded by the Belgian IAP BelVir.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Andrade, M. A. & Bork, P. (1995). *Nature Genet.* **11**, 115–116.
Bono, F., Cook, A. G., Grünwald, M., Ebert, J. & Conti, E. (2010). *Mol. Cell*, **37**, 211–222.
Brass, A. L., Dykxhoorn, D. M., Benita, Y., Yan, N., Engelman, A., Xavier, R. J., Lieberman, J. & Elledge, S. J. (2008). *Science*, **319**, 921–926.
Chook, Y. M. & Süel, K. E. (2011). *Biochim. Biophys. Acta*, **1813**, 1593–1606.
Christ, F., Thys, W., De Rijck, J., Gijsbers, R., Albanese, A., Arosio, D., Emiliani, S., Rain, J. C., Benarous, R., Cereseto, A. & Debyser, Z. (2008). *Curr. Biol.* **18**, 1192–1202.
Cook, A., Bono, F., Jinek, M. & Conti, E. (2007). *Annu. Rev. Biochem.* **76**, 647–671.
D'Angelo, M. A. & Hetzer, M. W. (2008). *Trends Cell Biol.* **18**, 456–466.
De Houwer, S., Demeulemeester, J., Thys, W., Taltynov, O., Zmajkovicova, K., Christ, F. & Debyser, Z. (2012). *J. Biol. Chem.* **287**, 34059–34068.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Forwood, J. K., Lange, A., Zachariae, U., Marfori, M., Preast, C., Grubmüller, H., Stewart, M., Corbett, A. H. & Kobe, B. (2010). *Structure*, **18**, 1171–1183.
Hirschfield, G. M. *et al.* (2010). *Nature Genet.* **42**, 655–657.
Hoelz, A., Debler, E. W. & Blobel, G. (2011). *Annu. Rev. Biochem.* **80**, 613–643.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
König, R. *et al.* (2008). *Cell*, **135**, 49–60.
Krishnan, L., Matreyek, K. A., Oztop, I., Lee, K., Tipper, C. H., Li, X., Dar, M. J., Kewalramani, V. N. & Engelman, A. (2010). *J. Virol.* **84**, 397–406.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Lai, M.-C., Lin, R.-I., Huang, S.-Y., Tsai, C.-W. & Tarn, W.-Y. (2000). *J. Biol. Chem.* **275**, 7950–7957.
Larue, R., Gupta, K., Wuensch, C., Shkriabai, N., Kessl, J. J., Danhart, E., Feng, L., Taltynov, O., Christ, F., Van Duyne, G. D., Debyser, Z., Foster, M. P. & Kvaratskhelia, M. (2012). *J. Biol. Chem.* **287**, 34044–34058.
Lee, B. J., Cansizoglu, A. E., Süel, K. E., Louis, T. H., Zhang, Z. & Chook, Y. M. (2006). *Cell*, **126**, 543–558.
Lee, S. J., Matsuura, Y., Liu, S. M. & Stewart, M. (2005). *Nature (London)*, **435**, 693–696.
Maertens, G. N., Cook, N. J., Wang, W., Hare, S., Gupta, S. S., Öztop, I., Lee, K., Pye, V. E., Cosnefroy, O., Snijders, A. P., KewalRamani, V. N., Fassati, A., Engelman, A. & Cherepanov, P. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 2728–2733.
Melià, M. J., Kubota, A., Ortolano, S., Vílchez, J. J., Gámez, J., Tanji, K., Bonilla, E., Palenzuela, L., Fernández-Cadenas, I., Pristoupilová, A., García-Arumí, E., Andreu, A. L., Navarro, C., Hirano, M. & Martí, R. (2013). *Brain*, **136**, 1508–1517.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nilsson, J., Weis, K. & Kjems, J. (2002). *J. Mol. Biol.* **318**, 583–593.
Stewart, M. (2007). *Nature Rev. Mol. Cell Biol.* **8**, 195–208.
Stewart, M., Kent, H. M. & McCoy, A. J. (1998). *J. Mol. Biol.* **284**, 1517–1527.
Taltynov, O., Demeulemeester, J., Christ, F., De Houwer, S., Tsirkone, V. G., Gerard, M., Weeks, S. D., Strelkov, S. V. & Debyser, Z. (2013). *J. Biol. Chem.* **288**, 25603–25613.
Torella, A., Fanin, M., Mutarelli, M., Peterle, E., Del Vecchio Blanco, F., Rispoli, R., Savarese, M., Garofalo, A., Piluso, G., Morandi, L., Ricci, G., Siciliano, G., Angelini, C. & Nigro, V. (2013). *PLoS One*, **8**, e63536.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Vetter, I. R., Arndt, A., Kutay, U., Görlich, D. & Wittinghofer, A. (1999). *Cell*, **97**, 635–646.
Weeks, S. D., Drinker, M. & Loll, P. J. (2007). *Protein Expr. Purif.* **53**, 40–50.
Yun, C. Y., Velazquez-Dones, A. L., Lyman, S. K. & Fu, X.-D. (2003). *J. Biol. Chem.* **278**, 18050–18055.